SANDIA REPORT

SAND2015-11019
Unlimited Release
Printed December 2015

Guidance for the Design and Adoption of Analytic Tools

Alisa Bandlow

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@osti.gov
Online ordering:	http://www.osti.gov/scitech

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.gov
Online order:	http://www.ntis.gov/search

SAND2015-11019
Unlimited Release
Printed December 2015

Guidance for the Design and Adoption of Analytic Tools

Alisa Bandlow
Operations Research and Knowledge Systems
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS1188

Abstract

The goal is to make software developers aware of common issues that can impede the adoption of analytic tools. This paper provides a summary of guidelines, lessons learned and existing research to explain what is currently known about what analysts want and how to better understand what tools they do and don't need.

Contents

FIGURES

1. Introduction

The "over the fence" development model is not effective; developers implement and evaluate systems in their labs and then throw these systems "over the fence" to the presumably grateful analysts. As Phillip Huxtable laments in his essay for *Challenges in Computational Social Modeling and Simulation for National Security Decision Making*, "Everyone in the community knows [the over the fence] approach doesn't work, yet the vast majority of analytic capability projects are executed this way, and most are unlikely to transition in any way that makes use of their apparent potential" [1].

The goal of this white paper is to make software developers aware of a variety of common issues that can impede the adoption of analytic tools. This report provides a summary of guidelines, lessons learned and existing research to explain what is currently known about the needs of information analysts and to better understand what tools they do and don't need.

2. Design Guidelines

While information analysts deal with a variety of data formats, these guidelines focus on designing for generic data and text analysis.

2.1. Features to Support General Data Analysis

When working on information retrieval tasks for numerous projects, Ben Shneiderman found himself rediscovering the same visual design principle each time, which he calls the Visual Information Seeking Mantra: "Overview first, zoom and filter, then details-on-demand" [2]. To help guide researchers and the development of prototypes to support this mantra, he proposes a type by task taxonomy that breaks down information visualizations into a set of basic data types and supporting tasks. At a minimum, tools should support the following seven tasks to support exploration of the seven basic data types:

1. **Overview:** gain an overview of the entire collection.
2. **Zoom:** zoom in on items of interest.
3. **Filter:** filter out uninteresting items.
4. **Details-on-demand:** select an item or group and get details when needed.
5. **Relate:** view relationships among items.
6. **History:** keep a history of actions to support undo, replay, and progressive refinement.
7. **Extract:** allow extraction of sub-collections and of query parameters.

2.2. Features to Support Text Analysis

The following sections describe features that have been compiled from existing research on analysts' work practices when dealing with text documents [3] [4] [5] [6]. These are actions that analysts frequently performed during analytic tasks or features requested by the analysts.

2.2.1. Mimic activities common to hardcopy documents

1. **Print documents:** ability to print text documents to hardcopy. Some analysts prefer to read and markup hardcopies first and later transfer these markups to the digital copies.
2. **Highlight text:** ability to highlight text within a document.
3. **Tag or annotate text:** add a user-defined tag or annotation/note to text within a document.
4. **Tag documents:** add multiple, user-defined tags to a document.
5. **Annotate documents:** add a note to a document that can be read or referenced in lieu of reading the document content.

6. **Organize documents into folders:** ability to organize documents into folders and subfolders.
7. **Lay out and organize documents spatially:** ability to spatially lay out and organize documents and folders, like organizing files on a computer desktop.

2.2.2. Extract and organize relevant information

1. **Copy and paste text from documents into a text editor:** ability to create a new document that consists of snippets of text from other documents and user-created notes and metadata.
2. **Create spreadsheets to collect information:** ability to create a new spreadsheet that consists of snippets of text from other documents and user-created notes and metadata.
3. **Sort documents based on time, tags, topics, geography or other attributes:** ability to sort and group documents on a variety of attributes. This allows the user to quickly explore potential themes and similarities in the document set.

2.2.3. Visualize the emerging story/hypothesis behind the data

1. **Create graphs to show relationships:** drawing tool capability to allow the user to create graphs that represent entities and their relationships.
2. **Link documents to nodes on a graph:** ability to link a document to a node on a graph, which makes it easier to track and reference information provenance for later reports.
3. **Have multiple views of the same data that are linked:** ability to highlight/select a single or group of documents in one view and also see those same documents highlighted when a different view is selected. For example, the user selects a group on entities on a user-created relational graph and then wants to see where these entities appear on a geospatial map.
4. **Review history of their work:** this can be implemented as bookmarks or screen captures (like in Palantir) of prior work states. Reviewing the analytic work history is important because "interpretations need to be audited, justified, revisited in the light of new information, and tackled by multiple analysts" [4].

2.2.4. Give the analyst control of the analytic process

1. **Undo/modify any automated tasks:** for automated tasks such as tagging and categorization, allow users to undo the automated task and modify the results of the automated task (e.g. change the tag, edit a category name, modify the contents of a category). This allows the user to correct any perceived "errors" made by the tool. This also makes the tool more flexible in supporting multiple analytic styles and processes.
2. **Manually perform the same tasks as the automated tasks:** allow users to perform tasks themselves without the aid of the tool (e.g. manually create a new category, manually tag content). This makes the tool more flexible by allowing users to take over when necessary.

2.3. Metrics

How to measure the efficacy of an analytic tool remains an open research question. Typical usability test measures such as completion rate and time on task do not fully address the question of whether or not a tool helps analysts make better decisions faster. To address this issue, the following metrics were developed for ARDA's Novel Intelligence for Massive Data (NIMD) program in the research area of Human Information Interaction [7]. Many of these metrics are easily captured as a logging feature if built into the tool from the beginning. This means decisions about what metrics are important and are success indicators must also be made at the beginning of the development process. The domain of web analytics can provide algorithms, methods and insight into how to understand search logs in terms of search success and user search behaviors.

Metrics for Human Information Interaction

- Efficiency
    - Time/search
    - Time/document read
- Effort
    - Number of documents accessed
    - Number of documents read
    - Document growth rate
    - Document growth type (cut/paste vs. typing)
- Accuracy
    - Evidence used in analysis
    - Number of hypotheses considered
    - Average system rank of documents viewed
- Confidence
    - User confidence ratings of findings
- Answer/Report Quality
    - Quality of report
    - Ranking of report
- Cognitive workload
    - Cognitive workload ratings (NASA Task Load Index (TLX) questionnaire [8])

3. What Analysts Want: Lessons Learned from Existing Research

3.1. Analysts want a single site or tool that addresses most of their needs, not a suite of single-action tools.

Many analytic tools are built as stand-alone tools that do not integrate with tools and technologies currently used by analysts.

Jim Powlen of Logos Technologies shared that during his discussions with analysts, they are eager for more effective tools, but complain that they already have too many of them. He found that they really want "one-stop shopping – a suite that will help them consolidate the information that they need … They have too many single-action tools, and that isn't really helping them" [9].

Intellipedia is an online, collaborative information sharing tool used by the U.S. Intelligence Community. Like all tools, it has its proponents and critics. A 2008 study on the usage of Intellipedia found that all interviewees (which included both proponents and critics) agreed that it provides "a quick, centralized, easy-to-use source of information" [10]. Its "one stop shop" nature contributes to its active use since 2006 [11].

3.2. Analysts need to understand what a tool/algorithm is doing in order to trust the results.

Helping analysts understand algorithms based on complex mathematics is a difficult task, but they will not use tools or algorithms for which they don't understand the underlying logic. They will not base their analyses and recommendations (and thus their reputations) on results and evidence that they don't fully understand or trust [1]. If algorithm results do not match their mental models, then they will brush off those results as wrong.

Huxtable, who works with Department of Defense organizations, notes, "I have seen many sophisticated capabilities rejected by analysts because they had no basis to trust the tool or method's validity and usefulness for their analytic tasks" [1].

For Sandia National Laboratories' Networks Grand Challenge, a prototype was created to aid in text analysis [12]. A major feature was an algorithm that automatically clustered documents into categories. However, test users had great difficulty in understanding the algorithm and its underlying logic, "commenting that documents addressing similar topics were located in different categories; or that unrelated documents were clustered together. As one tester commented, 'I just wouldn't do it that way'" [3]. Due to this confusion, test users did not want to use the algorithm's categories and requested the ability to create their own categories.

3.3. Algorithms and automation can sometimes hinder analysts' work and efficiency.

Many algorithms and automation tasks are meant to help information analysts deal with their big data problem of too many reports and not enough time. However, developers may automate tasks that they deem as trivial and time-consuming, which are actually critically important to the analysts' sensemaking activities.

A clustering algorithm was created as part of the Networks Grand Challenge to save analysts time by making it easier for them to determine which subset of documents to investigate [12]. Test users did not like the auto-created categories and instead requested the ability to rename and reconfigure the categories. A follow on study to investigate how analysts categorize documents found that reading and categorizing are not trivial tasks. This was how the analysts acquired new information, created mental models, and further refined and revised their understanding of the problem. As McNamara and Orlando-Gay noted, "If the process of categorization supports learning, as we believe it does, then having analytic tools 'hand over' a pre-established set of categories may actually *undermine* analysts' comprehension of a data set" [3].

3.4. Analysis is a highly iterative and spontaneous process. Automated workflows will be too restrictive.

Pirolli and Card describe analysis as consisting of two major loops of activities: a foraging loop that involves finding and extracting information and a sense making loop that involves creating a mental model to best fit the evidence [13]. As shown in Figure 1, this is a highly iterative process. The sensemaking loop itself may produce new questions that send the information analyst back to the foraging loop to find additional information.

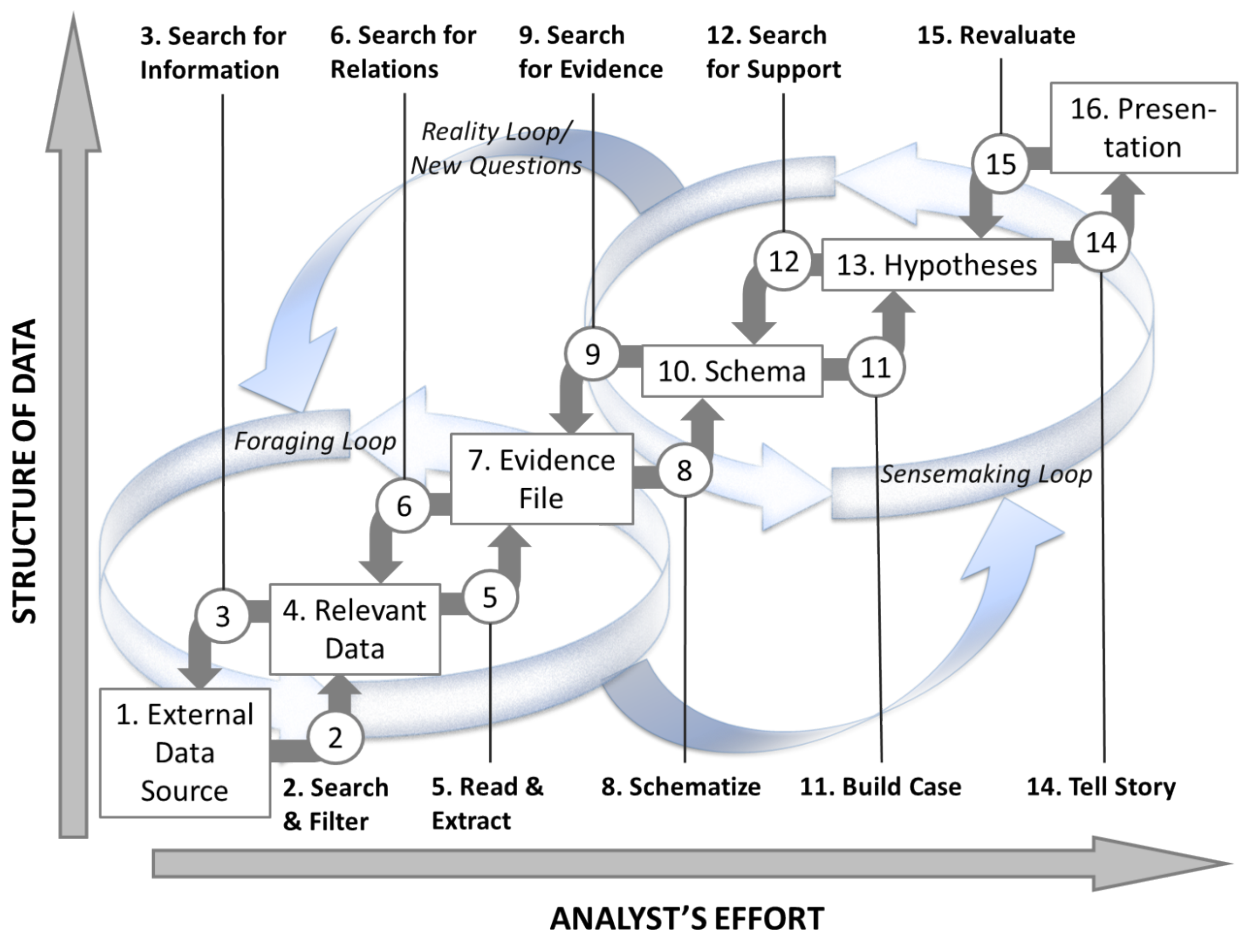

Figure 1. Notional model of sensemaking loop for analysis derived from Pirolli & Card's model [13].

In McNamara and Orlando-Gay's study of document categorization, they found that two-thirds of their participants exhibited similar behaviors: "reading, annotating, pausing, reaching for reports, shuffling through reports they had already read and annotated, then revising their categories" [3]. Andrews et. al. witnessed similar behavior [6].

In Chin's study of the analytical processes of analysts [5], his analysts "expressed that they would often abandon a systematic approach to satisfy time constraints … [and] may mentally conduct many aspects of an analysis without adhering to a particular investigative path". A tool with an automated workflow would be too restrictive for these analysts. The tool would probably be abandoned as soon as it conflicted with the constraints of their work environment.

3.5. The analytic process differs from analyst to analyst. Tools should focus on *supporting* analytic tasks instead of being a one size fits all solution.

Research shows that results for analytic tasks, such as grouping documents or hypotheses creation strategies, will vary from analyst to analyst [3] [5]. Much more research is needed to understand the breadth of analytic work practices and strategies employed. Thus, tools that focus on *supporting* analytic tasks instead of focusing on being the perfect, automated solution will have utility to a greater variety of analysts.

3.6. If the majority of a community can express reasonable opinions about a new technology, this typically indicates that it is on its way to widespread adoption. [14]

During their 2008 Intellipedia study, McNamara and Dixon were surprised that they did not encounter a single Defense Intelligence Agency staffer who had not heard of Intellipedia [10]. Most people were able to express reasonable opinions (both positive and negative) about how it worked and its impact on their workplace. Intellipedia is still actively in use [11].

4. Barriers to Technology Adoption

The usability and utility of a tool are not enough guarantee adoption, although they can be major factors. Usability guidelines and user testing are often focused on tool usage at the individual level. User testing typically occurs in an environment where the end user solely uses that tool to accomplish tasks. However, in practice, the tool will be used in a much richer and complex environment (Figure 2). The individual analyst is actually part of a larger work group with established work practices and protocols. The single tool will have to be used in concert with existing technologies and will need to be compatible with or integrated with them.

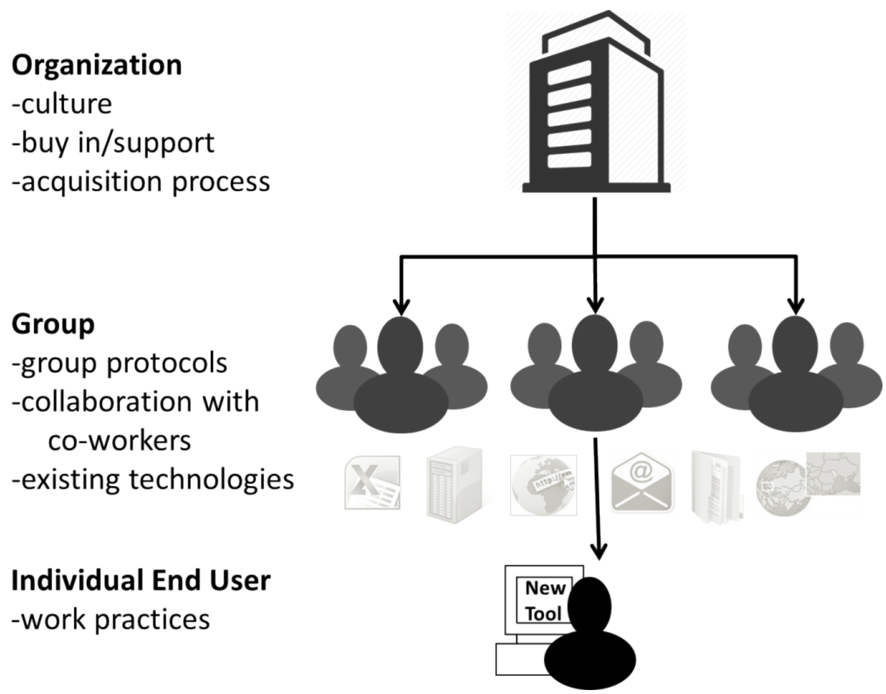

Figure 2. The analytic tool, while used by individual analysts, will need to operate in a complex environment.

4.1. The analysts' management needs to be clear about their expectations for a tool and how it fits into work practices.

In a study of Intellipedia usage at the Defense Intelligence Agency (DIA), McNamara and Dixon [10] found that analysts were uncertain about how Intellipedia fit into their work cycle. While the DIA office and division chiefs were encouraging usage of Intellipedia, they were at the same time expressing concerns that it was a work distraction.

4.2. If the goal of the tool does not match the work goals of the analysts, then the analysts have no motivation to use the tool.

Analysts' work goals are aligned with their performance assessments. Analysts are not motivated to use tools that do help them achieve their work goals because there is a lack of incentive, and there is often no penalty for not using the tool [15].

If the purpose of the tool is to change work practices or to accomplish new business goals, then the organizational structure needs to be modified to reward/penalize users for achieving/missing the tool's goal.

A tool can also inadvertently become a disincentive to the users if it makes it harder for them to achieve their main work goals. As Markus and Keil note, "[i]t is a well-known phenomenon that optimizing a subprocess can often lead to suboptimizing the larger process" [15]. This can occur when there is goal mismatch between the tool and the user: the user considers the tool's main goal to be a subprocess of his/her main work goal.

4.3. The organization often decides which technologies are used, not the analysts.

Many analysts do not have control over what is installed on their systems. Even if they want to use a new technology or tool, their organization ultimately decides what can be used. Thus, it is important to learn about the customer's acquisition process. It may be paramount to integrate with existing technologies to increase the chances of passing the acquisition process.

<u>Acquisition process:</u> Learn about the acquisition requirements in advance so that there is time and money to address post-development activities. The process could include Verification, Validation and Accreditation (VV&A) requirements, security scans that reveal issues that must be resolved, and rules about the use of third party products which require expensive last minute changes.

<u>Cost:</u> If a new tool requires the introduction of additional new technologies, the organization cannot or may not be willing to pay for them, annual licenses and additional maintenance.

<u>Cascading integration issues:</u> If one of these new technologies is meant to replace an existing technology, organization may be resistant to swap technologies even if the new one is promised to be better. First, users may like the existing technology and be resistant to change. Second, existing technologies are usually firmly integrated with other technologies, so swapping it out isn't a simple process. Multiple tools and systems will have to be reconfigured or rebuilt to integrate with the new technology, which adds to the cost. Third, other groups at the organization may also be reliant on the existing technology, which sets up a scenario where the organization will have to support multiple technologies that essentially perform the same function (again, at increased cost).

4.4. To overcome users' resistance to change, the new tool needs to provide a ten times improvement over their existing technology.

It can be difficult to convince users to give up their existing technologies even if a new technology is objectively better. According to Gourville, "consumers overvalue the existing benefits of an entrenched product by a factor of three, while developers overvalue the new benefits of their innovation by a factor of three. The result is a mismatch of nine to one, or 9 times, between what innovators think consumers desire and what consumers really want" [16]. This is due to the *endowment effect*, where people value products that they already possess more than those that they don't have; and the *status quo bias*, where giving up a possession feels like a loss and reduces the desire to trade up, even if there is a better alternative. The status quo bias also gets stronger with time: the users' reluctance to change technologies increases the longer they have been using it. Therefore in order to overcome people's natural tendency to see changing tools as a potential loss, the users have to view the relative benefit of the new tool as a 10 times improvement over their existing tool or method.

4.5. The developers and analyst managers must both be held responsible for the development of useful and valuable tools that are actually used. They must also have similar goals and incentives.

When the system isn't used, the analyst managers blame the system itself: if the system was good enough, then people would be using them without pressure from management. Since the system is often technically-sound, the developers redirect the blame back at the analysts: they built what the analysts said they wanted. These situations occur most frequently when the "true client" (the funder) is not the analyst manager, but some third party [15].

Markus and Keil believe these situations occur because developers, managers and their staff almost always have different objectives and incentives [15]. Developers are rewarded for delivering systems on time and on budget. They assume that management will take responsibility for getting staff to use the systems. The developers are not held accountable for creating systems that *are* used. Since the management of the intended users often has little control over these developers and the systems selected for implementation, they do not have an incentive to ensure that these systems are used and have business value.

In order to ensure the development of useful and valuable tools, both developers and management must be held responsible. One approach is concurrent system development, which requires (as a condition of funding approval) that every tool development project has a business/management sponsor who is responsible for achieving defined business goal [15]. This approach only works if the sponsor is evaluated and rewarded based on the achievement of these goals. The developers should also be rewarded by the organization based on how successfully the system is used. This ensures goal alignment between the project sponsors and developers.

A weakness of this approach is that sponsors can move to a different position. One mitigation strategy is to allow only projects that can produce significant results in two years or less. Another strategy is to review projects that have not produced results past the two year mark, especially those that have lost their initial sponsors.

5. Conclusion

Software developers focus on creating and implementing novel algorithms, models and systems. They do not always consider the bigger picture of successful technology adoption, which can include understanding the analytic activities a system should support, organization acquisition processes, and workplace incentives and goals.

The main theme is improved communication among developers, information analysts and management. Developers must collaborate with analysts to understand the analysts' work process, work environment and concerns. This will help developers create new tools that support the analysts' work rather than hinder them. Management on both sides must align their business objectives and incentives to ensure both sides gain from the project. This will aid in the development of useful tools with business value.

By providing a quick overview of common issues that can impede adoption, the goal is to raise developers' awareness so that these issues can be addressed early on and improve the chances of technology adoption.

Works Cited

[1] P. Huxtable, "Leveraging Computational Social Science for National Security," in *Challenges in Computation Social Modeling and Simulation for National Security Decision Making*, Ft. Belvoir, VA, Defense Threat Reduction Agency, Advanced Systems and Concepts Office, 2011, pp. 157-168.

[2] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proceedings of IEEE Symposium on Visual Languages*, Boulder, CO, 1996.

[3] L. McNamara and N. Orlando-Gay, "Reading, Sorting, Stacking, Shuffling: Low-Tech Studies Can Enhance Evaluation," in *Proceedings of the BELIV Workshop*, Seattle, WA, 2012.

[4] R. Walker, A. Slingsby, J. Dykes, K. Xu, J. Wood, P. Nguyen, D. Stephes, W. Wong and Y. Zheng, "An Extensible Framework for Provenance in Human Terrain Visual Analytics," *IEEE Transactions on Visualization and Computer Graphics,* vol. 19, no. 12, pp. 2139-2148, 2013.

[5] G. J. Chin, O. A. Kuchar and K. E. Wolf, "Exploring the Analytical Processes of Intelligence Analysts," in *Proceedings of ACM Computer Human Interaction*, Boston, MA, 2009.

[6] C. Andrews, A. Endert and C. North, "Space to Think: Large, High-Resolution Displays for Sensemaking," in *Proceedings of ACM Computer Human Interaction*, Atlanta, GA, 2010.

[7] E. Morse, M. P. Steves and J. Scholtz, "Metrics and Methodologies for Evaluating Technologies for Intelligence Analysts," in *Proceedings of Conference on Intelligence Analysis*, 2005.

[8] S. G. Hart and L. E. Staveland, "Development of a NASA-TLX (Task load index): Results of empirical and theoretical research," in *Human Mental Workload*, Amsterdam, North-Holland, 1988, pp. 139-183.

[9] R. Pool, *Field Evaluation in the Intelligence and Counterintelligence Context: Workshop Summary,* National Academy of Sciences, 2010.

[10] L. A. McNamara and N. A. Dixon, "Our Experience with Intellipedia: An Ethnographic Study at the Defense Intelligence Agency. SAND 2008-2187," Sandia National Laboratories, Albuquerque, NM, March 2008.

[11] J. Smathers, "FOIA Request for Intellipedia Usage Statistics," MuckRock.com, 9 January 2014. [Online]. Available: https://www.muckrock.com/foi/united-states-of-america-10/intellipedia-usage-statistics-10058/. [Accessed 15 September 2015].

[12] P. Kegelmeyer, "Network Discovery, Characterization, and Prediction: A Grand Challenge LDRD Final Report. SAND2010-8715," Sandia National Laboratories, Albuquerque, NM, 2010.

[13] P. Pirolli and S. Card, "The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis," in *Proceedings of International Conference on Intelligence Analysis*, 2005.

[14] E. Rogers, Diffusion of Innovations (Fifth Edition), New York: Basic Books, 2003.

[15] M. L. Markus and M. Keil, "If We Build It, They Will Come: Designing Information Systems That People Want to Use," *Sloan Management Review,* vol. 35, no. 4, pp. 11-25, Summer 1994.

[16] J. T. Gourville, "Eager Sellers and Stony Buyers: Understanding the Psychology of New-Product Adoption," *Harvard Business Review,* vol. 6, no. 84, pp. 98-106, 2006.

DISTRIBUTION

1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories